ESTIMATING MULTIVARIATE VARIANCE AND COVARIANCE COMPONENTS USING QUADRATIC AND BILINEAR FORMS*

BU-634-M　　　　by　　　　November, 1977

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

Abstract

Whenever variance components are estimated by quadratic forms, covariance components in the corresponding multivariate situation can be estimated by bilinear forms using the same matrices. Sampling variances and covariances of these estimators are derived.

1. The Univariate Case

1.1. The model

Notation first used by Hartley and Rao [1967] and many others can be adapted so as to write the general linear model for a vector $\underset{\sim}{x}$ of N realized values of the random variable X as

$$\underset{\sim}{x} = \mu \underset{\sim}{1} + \sum_{i=1}^{c} \underset{\sim}{Z}_i \underset{\sim}{b}_i \tag{1}$$

where μ is a general mean, $\underset{\sim}{1}$ is a vector of N ones, the $\underset{\sim}{b}_i$ for $i = 1, \cdots, c - 1$ are vectors of q_i effects and the $\underset{\sim}{Z}_i$ are known matrices, with $\underset{\sim}{b}_c = \underset{\sim}{e}$, a vector of

---

* Contribution to discussion of "An invariance property for first and second order moments of estimated variance-covariance components", by H. Ahrens, at the 19th "Session on Stochastics", Akademie der Wissenschaften, East Berlin, DDR, January 12, 1978.

Paper No. BU-634-M in the Biometrics Unit Mimeo Series.

residual error terms, and $\underset{\sim}{Z}_c = \underset{\sim}{I}_N$. For the random effects model (the model used in the Ahrens' paper, namely Eisenhart's Model II),

$$E(\underset{\sim}{b}_i) = \underset{\sim}{0}, \quad var(\underset{\sim}{b}_i) = \sigma_i^2 \underset{\sim}{I}_{q_i} \quad \text{and} \quad cov(\underset{\sim}{b}_i, \underset{\sim}{b}'_{i'}) = \underset{\sim}{0} \quad i \neq i' \tag{2}$$

with $q_c = N$ and $\sigma_c^2 = \sigma_e^2$ corresponding to $\underset{\sim}{b}_c = \underset{\sim}{e}$. From (1) and (2) the variance-covariance matrix of $\underset{\sim}{x}$, to be denoted by $\underset{\sim}{V}_x$, is

$$\underset{\sim}{V}_x = var(\underset{\sim}{x}) = \sum_{i=1}^{c} \sigma_i^2 \underset{\sim}{Z}_i \underset{\sim}{Z}'_i \ . \tag{3}$$

For mixed models, consisting of a mixture of fixed effects and random effects, one or more of the $\underset{\sim}{b}_i$'s (excluding $\underset{\sim}{b}_c$) represent fixed effects in which case, for those $\underset{\sim}{b}_i$'s, $E(\underset{\sim}{b}_i) \equiv \underset{\sim}{b}_i$ and $\sigma_i^2 \equiv 0$.

1.2. Estimation

We confine attention, as does Ahrens, to the class of unbiased estimators of σ_i^2 that are quadratic forms in $\underset{\sim}{x}$. Henderson's methods 1 (for random models), 2 (for mixed models with no interactions between fixed and random effects), and 3, and also Rao's MINQUE method, are in this class but, for example, the iterative methods of ML, REML, MIVQUE and iterative MINQUE are not. Within this class suppose that σ_i^2 is estimated by $\underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x}$; i.e.,

$$\hat{\sigma}_i^2 = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x} \ . \tag{4}$$

In most procedures for estimating σ_i^2 in the manner of (4), the quadratic form $\underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x}$ is not derived directly but is the consequence of taking the expected values of c other quadratic forms $\underset{\sim}{x}'\underset{\sim}{B}_i\underset{\sim}{x}$ (often sums of squares or analogous expressions). Then

$$E(\underset{\sim}{x}'\underset{\sim}{B}_i\underset{\sim}{x}) = tr(\underset{\sim}{B}_i\underset{\sim}{V}_x) + E(\underset{\sim}{x}')\underset{\sim}{B}_i E(\underset{\sim}{x}) \ . \tag{5}$$

In random models, $E(\underset{\sim}{x}) = \mu\underset{\sim}{1}$ and $\underset{\sim}{B}_i$ is usually such that $\underset{\sim}{1}'\underset{\sim}{B}_i = \underset{\sim}{0}$, so that the second term of (5) vanishes; and in mixed models, unbiased estimation of σ_i^2 by a quadratic form occurs only when $\underset{\sim}{B}$ and $E(\underset{\sim}{x})$ are such that the second term of (5) does vanish. Hence we can take (5) as

$$E(\underset{\sim}{x}'\underset{\sim}{B}_i\underset{\sim}{x}) = tr(\underset{\sim}{B}_i\underset{\sim}{V}_x) = tr(\underset{\sim}{B}_i \sum_{j=1}^{c} \sigma_j^2 \underset{\sim}{Z}_j\underset{\sim}{Z}_j')$$

$$= \sum_{j=1}^{c} \sigma_j^2 tr(\underset{\sim}{B}_i\underset{\sim}{Z}_j\underset{\sim}{Z}_j') = \sum_{j=1}^{c} k_{ij}\sigma_j^2 \ , \tag{6}$$

for k_{ij} defined as $k_{ij} = tr(\underset{\sim}{B}_i\underset{\sim}{Z}_j\underset{\sim}{Z}_j')$. Arraying each side of (6) for $i = 1, \cdots, c$ as a vector, the right-hand side would be $\underset{\sim}{K}\underset{\sim}{\sigma}^2$ for $\underset{\sim}{\sigma}^2 = [\sigma_1^2 \ \sigma_2^2 \ \cdots \ \sigma_c^2]'$. Then, on taking the inverse of $\underset{\sim}{K} = \{k_{ij}\}$ as $\underset{\sim}{K}^{-1} = \{k^{ij}\}$, and equating expected values to observed, we have

$$\hat{\sigma}_i^2 = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x} = \sum_{j=1}^{c} k^{ij}\underset{\sim}{x}'\underset{\sim}{B}_j\underset{\sim}{x} = \underset{\sim}{x}'\left(\sum_{j=1}^{c} k^{ij}\underset{\sim}{B}_j\right)\underset{\sim}{x} \tag{7}$$

so that in (4)

$$\underset{\sim}{A}_i = \sum_{j=1}^{c} k^{ij}\underset{\sim}{B}_j \ . \tag{8}$$

1.3. Example

The preceding description is nothing more than a formalization of the familiar procedure of equating expected values of quadratic forms to their observed values, as used in Henderson's Method 1, for example. There, for the 1-way classification model, $x_{ij} = \mu + \alpha_i + e_{ij}$ with $i = 1, \cdots, a$, and $j = 1, \cdots, n_i$ and $N = n_{\cdot} = \sum_{i=1}^{a} n_i$,

$$\underset{\sim}{x}'\underset{\sim}{B}_1\underset{\sim}{x} = \sum_i n_i(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 \quad \text{with} \quad \underset{\sim}{B}_1 = \sum_{i=1}^{a}{}^{+} \frac{1}{n_i}\underset{\sim}{J}_{n_i} - \frac{1}{N}\underset{\sim}{J}_N$$

and

$$\underset{\sim}{x}'\underset{\sim}{B}_2\underset{\sim}{x} = \sum_i\sum_j (\bar{x}_{ij} - \bar{x}_{i\cdot})^2 \quad \text{with} \quad \underset{\sim}{B}_2 = \underset{\sim}{I}_N - \sum_{i=1}^{a}{}^{+} \frac{1}{n_i}\underset{\sim}{J}_{n_i}$$

where $\underset{\sim}{J}_n$ is a square matrix of order n, with all elements unity, $\underset{\sim}{I}_n$ is an identity matrix of order n, and Σ^+ denotes a Kronecker sum of matrices. These are the $\underset{\sim}{B}_j$'s for (7) and (8), with j = 1 and 2. Then

$$E(\underset{\sim}{x}'\underset{\sim}{B}_1\underset{\sim}{x}) = \theta\sigma_\alpha^2 + (a-1)\sigma_e^2$$

and

$$E(\underset{\sim}{x}'\underset{\sim}{B}_2\underset{\sim}{x}) = (N-a)\sigma_e^2$$

for $\theta = N - \Sigma n_i^2/N$. Hence

$$\underset{\sim}{K} = \begin{bmatrix} \theta & a-1 \\ 0 & N-a \end{bmatrix} \text{ with } \underset{\sim}{K}^{-1} = \begin{bmatrix} \dfrac{1}{\theta} & \dfrac{-(a-1)}{(N-a)\theta} \\ 0 & \dfrac{1}{N-a} \end{bmatrix}$$

so giving the k^{ij}'s for (7) and (8).

1.4. Sampling variance

The sampling variance of the estimator in (4) based on normality assumptions is

$$v(\hat{\sigma}_i^2) = 2tr(\underset{\sim}{A}_i\underset{\sim}{V}_x)^2 \tag{9}$$

on using $\underset{\sim}{1}'\underset{\sim}{B}_i = \underset{\sim}{0}$ and hence $\underset{\sim}{1}'\underset{\sim}{A}_i = \underset{\sim}{0}$. From (3) we then get

$$v(\hat{\sigma}_i^2) = 2tr\left(\underset{\sim}{A}_i \sum_{j=1}^{c} \sigma_j^2 \underset{\sim}{Z}_j\underset{\sim}{Z}_j'\right)^2 \tag{10}$$

$$= 2\left[\sum_{j=1}^{c} \sigma_j^4 tr(\underset{\sim}{A}_i\underset{\sim}{Z}_j\underset{\sim}{Z}_j')^2 + 2\sum_{j=1}^{c}\sum_{j'>j}^{c} \sigma_j^2\sigma_{j'}^2 tr(\underset{\sim}{A}_i\underset{\sim}{Z}_j\underset{\sim}{Z}_j'\underset{\sim}{A}_i\underset{\sim}{Z}_{j'}\underset{\sim}{Z}_{j'}')\right] \tag{11}$$

Notation In (11) write

$$\lambda_{ii,jj} = \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j')^2 \quad \text{and} \quad \lambda_{ii,jj'} = \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j' \underset{\sim}{A}_i \underset{\sim}{Z}_{j'} \underset{\sim}{Z}_{j'}') \tag{12}$$

so that (11) becomes

$$v(\hat{\sigma}_i^2) = 2\left(\sum_{j=1}^{c} \sigma_j^4 \lambda_{ii,jj} + 2 \sum_{j=1}^{c} \sum_{j'>j}^{c} \sigma_j^2 \sigma_{j'}^2 \lambda_{ii,jj'} \right) . \tag{13}$$

More generally, define

$$\lambda_{ii',jj'} = \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j' \underset{\sim}{A}_{i'} \underset{\sim}{Z}_{j'} \underset{\sim}{Z}_{j'}') \tag{14}$$

so that both expressions in (12) are special cases of (14).

2. The Bivariate Case

2.1. Variance components

Suppose that on a random variable Y the vector of observations made on exactly the same set of observational units as yielded $\underset{\sim}{x}$ is $\underset{\sim}{y}$. Corresponding to (1), (2) and (3) we define a linear model for $\underset{\sim}{y}$ as

$$\underset{\sim}{y} = \mu^* \underset{\sim}{1} + \sum_{i=1}^{c} \underset{\sim}{Z}_i \underset{\sim}{b}_i^* \tag{15}$$

with

$$E(\underset{\sim}{b}_i^*) = \underset{\sim}{0}, \quad \mathrm{var}(\underset{\sim}{b}_i^*) = \sigma_i^{*2} \underset{\sim}{I}_{q_i}, \quad \text{and} \quad \mathrm{cov}(\underset{\sim}{b}_i^*, \underset{\sim}{b}_{i'}') = \underset{\sim}{0} \quad i \neq i' \tag{16}$$

and

$$\underset{\sim}{V}_y = \mathrm{var}(\underset{\sim}{y}) = \sum_{i=1}^{c} \sigma_i^{*2} \underset{\sim}{Z}_i \underset{\sim}{Z}_i' . \tag{17}$$

μ^*, $\underset{\sim}{b}_i^*$ and σ_i^{*2} play the same roles for $\underset{\sim}{y}$ as do μ, $\underset{\sim}{b}_i$ and σ_i^2 for $\underset{\sim}{x}$, and $\underset{\sim}{Z}_i$ is the same as for $\underset{\sim}{x}$ due to the observational units that yield $\underset{\sim}{y}$ being the same as those that yielded $\underset{\sim}{x}$. Hence equations (4) through (11) are true with $\underset{\sim}{y}$ and σ_i^{*2} replacing $\underset{\sim}{x}$ and σ_i^2, respectively. In particular, based on (7) and (13)

$$\hat{\sigma}_i^{*2} = \underset{\sim}{y}'\underset{\sim}{A}_i\underset{\sim}{y} = \sum_{j=1}^{c} k^{ij}\underset{\sim}{y}'\underset{\sim}{B}_j\underset{\sim}{y} = \underset{\sim}{y}'\left(\sum_{j=1}^{c} k^{ij}\underset{\sim}{B}_j\right)\underset{\sim}{y} \tag{18}$$

and

$$v(\hat{\sigma}_i^{*2}) = 2\left(\sum_{j=1}^{c} \sigma_j^{*4}\lambda_{ii,jj} + 2\sum_{j=1}^{c}\sum_{j'>j}^{c} \sigma_j^{*2}\sigma_{j'}^{*2}\lambda_{ii,jj'}\right) \tag{19}$$

2.2. Covariance components

Components of the covariance between $\underset{\sim}{x}$ and $\underset{\sim}{y}$ can be treated in exactly the same way. Define γ_i as the covariance between the k^{th} elements b_{ik} and b_{ik}^* of $\underset{\sim}{b}_i$ and $\underset{\sim}{b}_i^*$, respectively, for $k = 1, \cdots, q_i$; i.e.,

$$\gamma_i = \text{cov}(b_{ik}b_{ik}^*) = E(b_{ik}b_{ik}^*) , \tag{20}$$

the latter equality arising from (2) and (16). But note that

$$\text{cov}(b_{ik}b_{ik'}^*) = 0 \quad \text{for } k \neq k'$$

and (21)

$$\text{cov}(b_{ik}b_{i'k'}^*) = 0 \quad \text{for } i \neq i' \text{ and all } k \text{ and } k' .$$

As a result of (20) and (21), the matrix of covariances $\underset{\sim}{C}_{xy}$ between $\underset{\sim}{x}$ and $\underset{\sim}{y}$ is

$$\underset{\sim}{C}_{xy} = \text{cov}(\underset{\sim}{x}\underset{\sim}{y}') = \text{cov}\left(\sum_{i=1}^{c} \underset{\sim}{Z}_i\underset{\sim}{b}_i,\ \sum_{i=1}^{c} \underset{\sim}{b}_i^{*\prime}\underset{\sim}{Z}_i'\right) = \sum_{i=1}^{c} \gamma_i\underset{\sim}{Z}_i\underset{\sim}{Z}_i' , \tag{22}$$

similar to $\underset{\sim}{V}_x$ and $\underset{\sim}{V}_y$ of (3) and (17).

The estimator of γ_i corresponding to $\hat{\sigma}_i^2 = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x}$ and $\hat{\sigma}_i^{*2} = \underset{\sim}{y}'\underset{\sim}{A}_i\underset{\sim}{y}$ is $\hat{\gamma}_i = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{y}$ which is, as shown in Searle and Rounsaville [1974],

$$\hat{\gamma}_i = \mathbf{x}'\mathbf{A}_i\mathbf{y} = \tfrac{1}{2}[(\mathbf{x} + \mathbf{y})'\mathbf{A}_i(\mathbf{x} + \mathbf{y}) - \mathbf{x}'\mathbf{A}_i\mathbf{x} - \mathbf{y}'\mathbf{A}_i\mathbf{y}] = \tfrac{1}{2}(\hat{\sigma}_i^{**2} - \hat{\sigma}_i^2 - \hat{\sigma}_i^{*2}) \tag{23}$$

where $\hat{\sigma}_i^{**2}$ is the estimated variance component of the random variable $(X + Y)$ corresponding to $\hat{\sigma}_i^2$ and $\hat{\sigma}_i^{*2}$ of X and Y, respectively. Hence from (7)

$$\hat{\gamma}_i = \sum_{j=1}^{c} k^{ij}\mathbf{x}'\mathbf{B}_j\mathbf{y} \tag{24}$$

$$= \tfrac{1}{2}\sum_{j=1}^{c} k^{ij}[(\mathbf{x} + \mathbf{y})'\mathbf{B}_j(\mathbf{x} + \mathbf{y}) - \mathbf{x}'\mathbf{B}_j\mathbf{x} - \mathbf{y}'\mathbf{B}_j\mathbf{y}] .$$

2.3. Sampling variance of estimated covariance components

Consider two bilinear forms $\mathbf{x}_1'\mathbf{A}_{12}\mathbf{x}_2$ and $\mathbf{x}_3'\mathbf{A}_{34}\mathbf{x}_4$ where $\mathbf{A}_{12}$ and $\mathbf{A}_{34}$ are symmetric with $\mathbf{1}'\mathbf{A}_{12} = \mathbf{0}$ and $\mathbf{1}'\mathbf{A}_{34} = \mathbf{0}$, and $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$ have a multivariate normal distribution with means $\mu_r\mathbf{1}$ and covariances $\text{cov}(\mathbf{x}_r, \mathbf{x}_s') = \mathbf{C}_{rs}$ for $r, s = 1, \cdots, 4$. Then the covariance of these two bilinear forms is given by Searle [1971, Chapter 2, equation (58)] as

$$\text{cov}(\mathbf{x}_1'\mathbf{A}_{12}\mathbf{x}_2, \mathbf{x}_3'\mathbf{A}_{34}\mathbf{x}_4) = \text{tr}(\mathbf{A}_{12}\mathbf{C}_{23}\mathbf{A}_{34}\mathbf{C}_{41}) + \text{tr}(\mathbf{A}_{12}\mathbf{C}_{24}\mathbf{A}_{43}\mathbf{C}_{31}) . \tag{25}$$

We use this result to derive first the sampling variance of $\hat{\gamma}_i = \mathbf{x}'\mathbf{A}_i\mathbf{y}$ and then the sampling covariances of the estimated components of variance and covariance.

In (25) put $\mathbf{x}_1 = \mathbf{x}_3 = \mathbf{x}$, $\mathbf{x}_2 = \mathbf{x}_4 = \mathbf{y}$ and $\mathbf{A}_{12} = \mathbf{A}_{34} = \mathbf{A}_i$, so that $\mathbf{C}_{23} = \mathbf{C}_{41} = \mathbf{C}_{xy}$ of (20) and $\mathbf{C}_{31} = \mathbf{V}_x$ and $\mathbf{C}_{42} = \mathbf{V}_y$. Then (25) gives

$$v(\hat{\gamma}_i) = \text{tr}(\mathbf{A}_i\mathbf{C}_{xy})^2 + \text{tr}(\mathbf{A}_i\mathbf{V}_x\mathbf{A}_i\mathbf{V}_y) . \tag{26}$$

By the nature of $\mathbf{V}_x$, $\mathbf{V}_y$ and $\mathbf{C}_{xy}$ in (3), (17) and (22), this is

$$v(\hat{\gamma}_i) = \mathrm{tr}\Big(\sum_{j=1}^{c} \gamma_i \underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j'\Big)^2 + \mathrm{tr}\Big(\sum_{j=1}^{c} \sigma_j^2 \underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j'\Big)\Big(\sum_{j=1}^{c} \sigma_j^{*2} \underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j'\Big)$$

$$= \Big[\sum_{j=1}^{c} \gamma_i^2 \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j')^2 + 2 \sum_{j=1}^{c} \sum_{j'>j}^{c} \gamma_j \gamma_{j'} \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j' \underset{\sim}{A}_i \underset{\sim}{Z}_{j'} \underset{\sim}{Z}_{j'}')\Big]$$

$$+ \Big[\sum_{j=1}^{c} \sigma_j^2 \sigma_j^{*2} \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j')^2 + \sum_{j=1}^{c} \sum_{j'>j}^{c} (\sigma_j^2 \sigma_{j'}^{*2} + \sigma_{j'}^2 \sigma_j^{*2}) \mathrm{tr}(\underset{\sim}{A}_i \underset{\sim}{Z}_j \underset{\sim}{Z}_j' \underset{\sim}{A}_i \underset{\sim}{Z}_{j'} \underset{\sim}{Z}_{j'}')\Big]$$

$$= \sum_{j=1}^{c} (\gamma_j^2 + \sigma_j^2 \sigma_j^{*2}) \lambda_{ii,jj} + \sum_{j=1}^{c} \sum_{j'>j}^{c} (2\gamma_j \gamma_{j'} + \sigma_j^2 \sigma_{j'}^{*2} + \sigma_{j'}^2 \sigma_j^{*2}) \lambda_{ii,jj'} \ . \qquad (27)$$

This is a generalization of result (14) given in Searle [1956] for the 1-way classification.

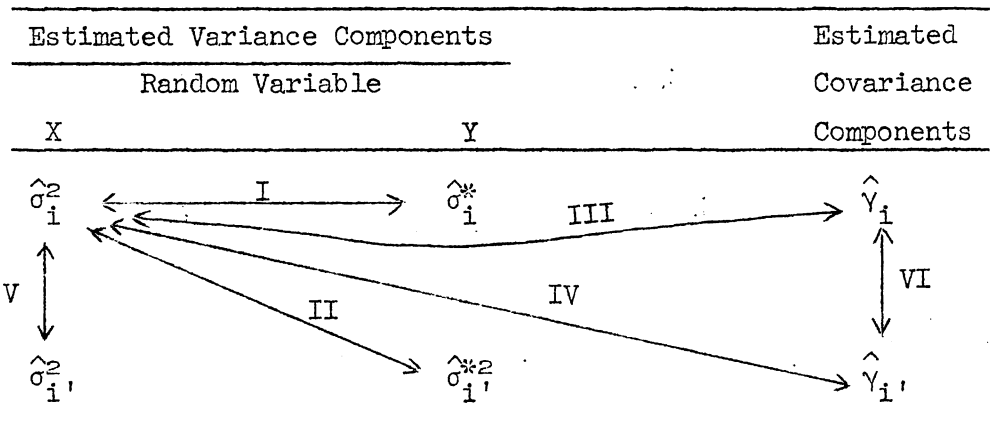
2.4. Similarities between covariance and variance components

The well-known invariance property (as Ahrens' paper calls it) of the first order moments is apparent in equations (7), (18) and (24), namely that $\hat{\sigma}_i^{*2}$ is the same linear function of the same quadratic forms in $\underset{\sim}{y}$ as $\hat{\sigma}_i^2$ is in $\underset{\sim}{x}$, and $\hat{\gamma}_i$ is the same linear function of bilinear forms in $\underset{\sim}{x}$ and $\underset{\sim}{y}$ using the same matrices as used in the quadratic forms. Furthermore, the same kind of invariance property of second-order moments, in this case sampling variances, – the moments considered in the Ahrens paper – is evident in equations (13), (19) and (27). In those equations the coefficients of $2\sigma_j^4$, $2\sigma_j^{*4}$ and $(\gamma_j^2 + \sigma_j^2\sigma_j^{*2})$ in $v(\hat{\sigma}_i^2)$, $v(\hat{\sigma}_i^{*2})$ and $v(\hat{\gamma}_i)$, respectively, are all the same, namely $\lambda_{ii,jj}$, and the coefficients of $4\sigma_j^2\sigma_{j'}^2$, $4\sigma_j^{*2}\sigma_{j'}^{*2}$ and $(2\gamma_j\gamma_{j'} + \sigma_j^2\sigma_{j'}^{*2} + \sigma_{j'}^2\sigma_j^{*2})$ are also all the same, $\lambda_{ii,jj'}$ for $j \neq j'$. And these results are true, not only for the special cases that Ahrens considers but, more widely, for all cases and estimation methods where variance components estimators are quadratic forms.

2.5. Sampling covariances

Equation (25) can also be used to derive sampling covariances between estimated components of variance and covariance in the bivariate (and multivariate) case. The arrowed lines in the following figure numbered I through VI identify the six different kinds of covariances.

Six Covariances



The Estimators

$\hat{\sigma}_i^2 = \mathbf{x}'\mathbf{A}_i\mathbf{x}$ $\qquad \hat{\sigma}_i^{*2} = \mathbf{y}'\mathbf{A}_i\mathbf{y}$ $\qquad \hat{\gamma}_i = \mathbf{x}'\mathbf{A}_i\mathbf{y}$

$\hat{\sigma}_{i'}^2 = \mathbf{x}'\mathbf{A}_{i'}\mathbf{x}$ $\qquad \hat{\sigma}_{i'}^{*2} = \mathbf{y}'\mathbf{A}_{i'}\mathbf{y}$ $\qquad \hat{\gamma}_{i'} = \mathbf{x}'\mathbf{A}_{i'}\mathbf{y}$

The manner in which (25) is used for each of the six covariances is shown in the following table:

Table 1

	Term in (25)									
	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{A}_{12}$	$\mathbf{A}_{34}$	$\mathbf{C}_{23}$	$\mathbf{C}_{41}$	$\mathbf{C}_{24}$	$\mathbf{C}_{31}$
I: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{x},\ \mathbf{y}'\mathbf{A}_i\mathbf{y})$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{y}$	$\mathbf{A}_i$	$\mathbf{A}_i$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$
II: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{x},\ \mathbf{y}'\mathbf{A}_{i'}\mathbf{y})$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{y}$	$\mathbf{A}_i$	$\mathbf{A}_{i'}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$
III: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{x},\ \mathbf{x}'\mathbf{A}_i\mathbf{y})$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{A}_i$	$\mathbf{A}_i$	$\mathbf{V}_x$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{V}_x$
IV: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{x},\ \mathbf{x}'\mathbf{A}_{i'}\mathbf{y})$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{A}_i$	$\mathbf{A}_{i'}$	$\mathbf{V}_x$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{V}_x$
V: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{x},\ \mathbf{x}'\mathbf{A}_{i'}\mathbf{x})$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{x}$	$\mathbf{A}_i$	$\mathbf{A}_{i'}$	$\mathbf{V}_x$	$\mathbf{V}_x$	$\mathbf{V}_x$	$\mathbf{V}_x$
VI: $\mathrm{cov}(\mathbf{x}'\mathbf{A}_i\mathbf{y},\ \mathbf{x}'\mathbf{A}_{i'}\mathbf{y})$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{x}$	$\mathbf{y}$	$\mathbf{A}_i$	$\mathbf{A}_{i'}$	$\mathbf{C}_{xy}$	$\mathbf{C}_{xy}$	$\mathbf{V}_y$	$\mathbf{V}_x$

Carrying out these special cases of (25) gives the six covariances as follows:

$$\text{I:}\quad \operatorname{cov}(\hat{\sigma}_i^2, \hat{\sigma}_i^{*2}) = 2\left(\sum_{j=1}^{c} \gamma_j^2\lambda_{ii,jj} + 2\sum_{j=1}^{c}\sum_{j'>j}^{c} \gamma_j\gamma_{j'}\lambda_{ii,jj'}\right)$$

$$\text{II:}\quad \operatorname{cov}(\hat{\sigma}_i^2, \hat{\sigma}_{i'}^{*2}) = 2\left(\sum_{j=1}^{c} \gamma_j^2\lambda_{ii',jj} + 2\sum_{j=1}^{c}\sum_{j'>j}^{c} \gamma_j\gamma_{j'}\lambda_{ii',jj'}\right)$$

$$\text{III:}\quad \operatorname{cov}(\hat{\sigma}_i^2, \hat{\gamma}_i) = 2\left[\sum_{j=1}^{c} \sigma_j^2\gamma_j\lambda_{ii,jj} + \sum_{j=1}^{c}\sum_{j'>j}^{c} (\sigma_j^2\gamma_{j'} + \sigma_{j'}^2\gamma_j)\lambda_{ii,jj'}\right]$$

$$\text{IV:}\quad \operatorname{cov}(\hat{\sigma}_i^2, \hat{\gamma}_{i'}) = 2\left[\sum_{j=1}^{c} \sigma_j^2\gamma_j\lambda_{ii',jj} + \sum_{j=1}^{c}\sum_{j'>j}^{c} (\sigma_j^2\gamma_{j'} + \sigma_{j'}^2\gamma_j)\lambda_{ii',jj'}\right]$$

$$\text{V:}\quad \operatorname{cov}(\hat{\sigma}_i^2, \hat{\sigma}_{i'}^2) = 2\left(\sum_{j=1}^{c} \sigma_j^4\lambda_{ii',jj} + 2\sum_{j=1}^{c}\sum_{j'>j}^{c} \sigma_j^2\sigma_{j'}^2\lambda_{ii',jj'}\right)$$

$$\text{VI:}\quad \operatorname{cov}(\hat{\gamma}_i, \hat{\gamma}_{i'}) = \sum_{j=1}^{c} (\gamma_j^2 + \sigma_j^2\sigma_j^{*2})\lambda_{ii',jj} + \sum_{j=1}^{c}\sum_{j'>j}^{c} (2\gamma_j\gamma_{j'} + \sigma_j^2\sigma_{j'}^{*2} + \sigma_{j'}^2\sigma_j^{*2})\lambda_{ii',jj'} \;.$$

The kind of similarities noted in Section 2.4 can be seen here too. In all cases the coefficients of the first and second terms are either $\lambda_{ii,jj}$ and $\lambda_{ii,jj'}$ respectively, or $\lambda_{ii',jj}$ and $\lambda_{ii',jj'}$ depending upon the subscripts of the terms for which the covariance is being stated: either i and i, or i and i' .

3. The General Multivariate Case

When three variables are involved, with data vectors $\underset{\sim}{x}$, $\underset{\sim}{y}$ and $\underset{\sim}{z}$ and variance-covariance matrix

$$\underset{\sim}{V} = \operatorname{var}\begin{bmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \\ \underset{\sim}{z} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{V}_x & \underset{\sim}{C}_{xy} & \underset{\sim}{C}_{xz} \\ \underset{\sim}{C}_{yx} & \underset{\sim}{V}_y & \underset{\sim}{C}_{yz} \\ \underset{\sim}{C}_{zx} & \underset{\sim}{C}_{zy} & \underset{\sim}{V}_z \end{bmatrix},$$

any quadratic form such as $\hat{\sigma}_i^2 = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x}$ can be written as

$$\hat{\sigma}_i^2 = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{x} = [\underset{\sim}{x}' \ \underset{\sim}{y}' \ \underset{\sim}{z}'] \begin{bmatrix} \underset{\sim}{A}_i & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \\ \underset{\sim}{z} \end{bmatrix} \tag{28}$$

so that its variance is

$$v(\hat{\sigma}_i^2) = 2\mathrm{tr}\left[\begin{pmatrix} \underset{\sim}{A}_i & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \end{pmatrix}\underset{\sim}{V}\right]^2 \tag{29}$$

$$= 2\mathrm{tr}(\underset{\sim}{A}_i\underset{\sim}{V}_x)^2$$

just as in (9). Similarly, any bilinear form such as $\hat{\gamma}_i = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{y}$ can be expressed as

$$\hat{\gamma}_i = \underset{\sim}{x}'\underset{\sim}{A}_i\underset{\sim}{y} = [\underset{\sim}{x}' \ \underset{\sim}{y}' \ \underset{\sim}{z}'] \ \tfrac{1}{2}\begin{bmatrix} \underset{\sim}{0} & \underset{\sim}{A}_i & \underset{\sim}{0} \\ \underset{\sim}{A}_i & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \\ \underset{\sim}{z} \end{bmatrix} \tag{30}$$

with variance

$$v(\hat{\gamma}_i) = 2\mathrm{tr}\left[\tfrac{1}{2}\begin{pmatrix} \underset{\sim}{0} & \underset{\sim}{A}_i & \underset{\sim}{0} \\ \underset{\sim}{A}_i & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \end{pmatrix}\underset{\sim}{V}\right]^2$$

$$= \tfrac{1}{2}\mathrm{tr}\begin{bmatrix} \underset{\sim}{A}_i\underset{\sim}{C}_{yx} & \underset{\sim}{A}_i\underset{\sim}{V}_y & \underset{\sim}{A}_i\underset{\sim}{V}_{yz} \\ \underset{\sim}{A}_i\underset{\sim}{V}_x & \underset{\sim}{A}_i\underset{\sim}{C}_{xy} & \underset{\sim}{A}_i\underset{\sim}{V}_{xz} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix}^2$$

$$= \tfrac{1}{2}\mathrm{tr}\left[(\underset{\sim}{A}_i\underset{\sim}{C}_{yx})^2 + \underset{\sim}{A}_i\underset{\sim}{V}_y\underset{\sim}{A}_i\underset{\sim}{V}_x + \underset{\sim}{A}_i\underset{\sim}{V}_x\underset{\sim}{A}_i\underset{\sim}{V}_y + (\underset{\sim}{A}_i\underset{\sim}{C}_{xy})^2\right]$$

$$= \mathrm{tr}(\underset{\sim}{A}_i\underset{\sim}{C}_{xy})^2 + \mathrm{tr}(\underset{\sim}{A}_i\underset{\sim}{V}_x\underset{\sim}{A}_i\underset{\sim}{V}_y) \tag{31}$$

as in (26).

The kinds of expression illustrated in (28) and (30) for a 3-variable situation carry over identically to the many-variable case. Similarly, results (29) and (31), showing that expressions derived for the 2-variable case also hold for 3 variables, extend quite naturally for any number of variables; i.e., the results of Section 2 hold quite generally for any pair of variables in a multivariate situation.

4. Example: The 1-way Classification

The model for the 1-way classification can be written as $x_{ij} = \mu + \alpha_i + e_{ij}$ with $i = 1, \cdots, a$, and $j = 1, \cdots, n_i$, with variance components σ^2_α and σ^2_e. Then, as indicated at the end of Section 1.3, the estimated components which we shall here also denote as functions $\hat{\sigma}^2_\alpha(\underset{\sim}{x})$ and $\hat{\sigma}^2_e(\underset{\sim}{y})$ are, for $\theta = N - \sum_{i=1}^{a} n_i^2/N$

$$\hat{\sigma}^2_\alpha = \hat{\sigma}^2_\alpha(\underset{\sim}{x}) = \left[\sum_{i=1}^{a} n_i(\bar{x}_{i.} - \bar{x}_{.})^2 - (a - 1)\sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{i.})^2/(N - a)\right]/\theta \tag{32}$$

and

$$\hat{\sigma}^2_e = \hat{\sigma}^2_e(\underset{\sim}{x}) = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{i.})^2/(N - a)\ . \tag{33}$$

These are examples of (7). Then, by (18)

$$\hat{\sigma}^{*2}_\alpha = \hat{\sigma}^2_\alpha(\underset{\sim}{y}) \quad \text{and} \quad \hat{\sigma}^{*2}_e = \hat{\sigma}^2_e(\underset{\sim}{y})$$

and by (23)

$$\hat{\gamma}_\alpha = \tfrac{1}{2}\left[\hat{\sigma}^2_\alpha(\underset{\sim}{x}+\underset{\sim}{y}) - \hat{\sigma}^2_\alpha(\underset{\sim}{x}) - \hat{\sigma}^2_\alpha(\underset{\sim}{y})\right] \quad \text{and} \quad \hat{\gamma}_e = \left[\hat{\sigma}^2_e(\underset{\sim}{x}+\underset{\sim}{y}) - \hat{\sigma}^2_e(\underset{\sim}{x}) - \hat{\sigma}^2_e(\underset{\sim}{y})\right].$$

From Searle [1971, p. 474], using the definitions

$$S_2 = \sum_{i=1}^{a} n_i^2 \quad \text{and} \quad S_3 = \sum_{i=1}^{a} n_i^3$$

the variances of, and covariance between, $\hat{\sigma}^2_\alpha$ and $\hat{\sigma}^2_e$ of (32) and (33) are as follows:

$$v(\hat{\sigma}^2_\alpha) = 2\sigma^4_\alpha \frac{N^2S_2 + S_2^2 - 2NS_3}{(N^2 - S_2^2)} + 2\sigma^4_e \frac{N^2(N-1)(a-1)}{(N-a)(N^2-S_2)^2} + 4\sigma^2_\alpha\sigma^2_e \frac{N}{N^2 - S_2} \tag{34}$$

$$v(\hat{\sigma}^2_e) = \frac{2\sigma^4_e}{N-a} \tag{35}$$

and

$$\text{cov}(\hat{\sigma}^2_\alpha, \hat{\sigma}^2_e) = 2\sigma^4_e \frac{-N(a-1)}{(N-a)(N^2 - S_2)} \tag{36}$$

Writing

$$\sigma^2_\alpha \equiv \sigma^2_1 \quad \text{and} \quad \sigma^2_e \equiv \sigma^2_2 \tag{37}$$

and comparing (34) and (35) with (13), and (36) with V below Table 1, gives the λ's as follows:

$$\lambda_{11,11} = \frac{N^2S_2 + S_2^2 - 2NS_3}{(N^2 - S_2)^2} \qquad \lambda_{11,22} = \frac{N^2(N-1)(a-1)}{(N-a)(N^2-S_2)^2} \qquad \lambda_{11,12} = \frac{N}{N^2 - S_2}$$

$$\lambda_{22,11} = 0 \qquad \lambda_{22,22} = \frac{1}{N - a} \qquad \lambda_{22,12} = 0$$

$$\lambda_{12,11} = 0 \qquad \lambda_{12,22} = \frac{-N(a-1)}{(N-a)(N^2-S_2)} \qquad \lambda_{12,12} = 0 .$$

Then the fifteen different covariances among $\hat{\sigma}^2_\alpha$, $\hat{\sigma}^2_e$, $\hat{\sigma}^{*2}_\alpha$, $\hat{\sigma}^{*2}_e$, $\hat{\gamma}_\alpha$ and $\hat{\gamma}_e$ are obtained from the six different kinds of covariances given below Table 1.

$$\text{I:}\quad \operatorname{cov}(\hat{\sigma}^2_\alpha\,,\,\hat{\sigma}^{*2}_\alpha) = 2(\lambda_{11,11}\gamma^2_\alpha + \lambda_{11,22}\gamma^2_e + 2\lambda_{11,12}\gamma_\alpha\gamma_e)$$

$$\operatorname{cov}(\hat{\sigma}^2_e\,,\,\hat{\sigma}^{*2}_e) = 2\lambda_{22,22}\gamma^2_e$$

$$\text{II:}\quad \operatorname{cov}(\hat{\sigma}^2_\alpha\,,\,\hat{\sigma}^{*2}_e) = 2\lambda_{12,22}\gamma^2_e = \operatorname{cov}(\hat{\sigma}^{*2}_\alpha,\,\hat{\sigma}^2_e)$$

$$\text{III:}\quad \operatorname{cov}(\hat{\sigma}^2_\alpha\,,\,\hat{\gamma}_\alpha) = 2\left[\lambda_{11,11}\sigma^2_\alpha\gamma_\alpha + \lambda_{11,22}\sigma^2_e\gamma_e + \lambda_{11,12}(\sigma^2_\alpha\gamma_e + \sigma^2_e\gamma_\alpha)\right]$$

$$\operatorname{cov}(\hat{\sigma}^{*2}_\alpha,\,\hat{\gamma}_\alpha) = 2\left[\lambda_{11,11}\sigma^{*2}_\alpha\gamma_\alpha + \lambda_{11,22}\sigma^{*2}_e\gamma_e + \lambda_{11,12}(\sigma^{*2}_\alpha\gamma_e + \sigma^{*2}_e\gamma_\alpha)\right]$$

$$\operatorname{cov}(\hat{\sigma}^2_e\,,\,\hat{\gamma}_e) = 2\lambda_{22,22}\sigma^2_e\gamma_e$$

$$\operatorname{cov}(\hat{\sigma}^{*2}_e,\,\hat{\gamma}_e) = 2\lambda_{22,22}\sigma^{*2}_e\gamma_e$$

$$\text{IV:}\quad \operatorname{cov}(\hat{\sigma}^2_\alpha\,,\,\hat{\gamma}_e) = 2\lambda_{12,22}\sigma^2_e\gamma_e = \operatorname{cov}(\hat{\sigma}^2_e\,,\,\hat{\gamma}_\alpha)$$

$$\operatorname{cov}(\hat{\sigma}^{*2}_\alpha,\,\hat{\gamma}_e) = 2\lambda_{12,22}\sigma^{*2}_e\gamma_e = \operatorname{cov}(\hat{\sigma}^{*2}_e,\,\hat{\gamma}_\alpha)$$

$$\text{V:}\quad \operatorname{cov}(\hat{\sigma}^2_\alpha\,,\,\hat{\sigma}^2_e) = 2\lambda_{12,22}\sigma^4_e\,, \qquad \text{as in (36)}$$

$$\operatorname{cov}(\hat{\sigma}^{*2}_\alpha,\,\hat{\sigma}^{*2}_e) = 2\lambda_{12,22}\sigma^{*4}_e$$

$$\text{VI:}\quad \operatorname{cov}(\hat{\gamma}_\alpha\,,\,\hat{\gamma}_e) = \lambda_{12,22}(\gamma^2_e + \sigma^2_e\sigma^{*2}_e)\ .$$

References

Hartley, H. O. and Rao, J. N. K. [1967]. Maximum likelihood estimation for the mixed analysis of variance model. Biometrika 54, 93-108.

Searle, S. R. [1956]. Matrix methods in variance and covariance components analysis. Ann. Math. Stat. 27, 737-748.

Searle, S. R. [1971]. Linear Models. Wiley, New York.

Searle, S. R. and Rounsaville, T. R. [1974]. A note on estimating covariance components. The American Statistician 28, No. 2, 67-68.